DOCUMENT RESUME

ED 051 313 TM 000 705

TITLE	Proceedings of the Conference on Educational Accountability.
INSTITUTION	Educational Testing Service, Princeton, N.J.
PUB DATE	Jun 71
NOTE	67p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Academic Achievement, Behavioral Objectives, *Conference Reports, Criterion Referenced Tests, *Educational Accountability, Educational Finance, Educational Improvement, Educational Objectives, Educational Planning, *Educational Quality, Evaluation, *Instructional Programs, *Performance Contracts, Performance Criteria, Program Development, Public Education, School Community Relationship, Teacher Responsibility, Teaching Quality

ABSTRACT

The philosophy, strategies, and pitfalls of educational accountability, and information on the development and implementation of accountability programs are outlined in this conference report. Educational accountability is a term used in connection with activities such as assessment, evaluation, auditing, and performance contacting. The speakers' papers include: "The Means and Ends of Accountability" (Erick L. Lindman); "Issues in Implementation I" (Mark R. Shedd); "Issues in Implementation II" (Francis Keppel); "Public Expectations" (James E. Allen, Jr.); "The Role of Evaluation" (Henry S. Dyer); and "The Future of Accountability" (Edythe J. Gaines). (AG)

ED051313

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Proceedings of the
Conference
on

Educational Accountability

CHICAGO, ILLINOIS

JUNE, 1971

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Educational Testing Service
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

TM 000 205

Sponsored by

EDUCATIONAL TESTING SERVICE

Copyright © 1971 by Educational Testing Service. All rights reserved.

ED051313

Conference

on

Educational Accountability

William W. Turnbull
CHAIRMAN

EDUCATIONAL TESTING SERVICE
Princeton, New Jersey
Berkeley, California
Evanston, Illinois

3

C O N T R I B U T O R S

James E. Allen, Jr.
Visiting Lecturer in Education and Public Affairs
Woodrow Wilson School, Princeton University
and
Special Advisor
Educational Testing Service

Henry S. Dyer
Vice President
Educational Testing Service

Edythe J. Gaines
Community Superintendent
New York City

Francis Keppel
Chairman of the Board
General Learning Corporation

Erick L. Lindman
Professor of Education
University of California at Los Angeles
and
Member, California Advisory Commission
on School District Budgeting and Accounting

Mark R. Shedd
Superintendent of Schools
Philadelphia

William W. Turnbull
President
Educational Testing Service

C O N T E N T S

	Section
Preface - William W. Turnbull	A
* * *	
The Means and Ends of Accountability	
Erick L. Lindman	B
Issues in Implementation	
Mark R. Shedd	C
Francis Keppel	D
Public Expectations	
James E. Allen, Jr.	E
The Role of Evaluation	
Henry S. Dyer	F
The Future of Accountability	
Edythe J. Gaines	G

PREFACE

William W. Turnbull

Both professional and public attention is focused these days on the concept of educational accountability. The word "accountability" has as many meanings as people care to give it and is often used in connection with such activities as assessment, evaluation, auditing, and performance contracting.

So great is the interest -- and so meager the clarification of the many issues and problems involved -- that the need for a comprehensive look at the concept prompted Educational Testing Service to sponsor a series of conferences on accountability.

We are indeed fortunate that some of the most knowledgeable and thoughtful people concerned with the philosophy, strategies, and pitfalls of accountability in education accepted our invitation to participate in this conference. Each of them has provided a substantive and challenging contribution to better understanding of what is involved in developing and implementing accountability programs of integrity and merit.

Because of the urgent need for dissemination of information about accountability the speakers' papers, in their pre-conference form, have been assembled in this booklet for immediate distribution.

THE MEANS AND ENDS OF ACCOUNTABILITY

Erick L. Lindman

Benefits and Costs of Education

A few years ago a hardy band of program budgeteers gathered together the elements of their planning, programming, and budgeting system, packed them in a waterproof box, and departed from the Pentagon where they had spent many fruitful years. Their departure was without ceremony. There seemed to be neither sadness nor joy at their leaving; but evidence of relief was clearly visible beneath the surface.

The sturdy band set sail for education land where they hoped to spread the gospel of PPBS. When they reached education land they were admitted without delay and were warmly welcomed, for they promised something for everyone. For school boards they promised more education for the school dollar. For administrators they promised a management information system with modern built-in feedback controls; and for teachers they promised that the budget would no longer be the mysterious obstacle to their plans, but instead would become a helpful genie.

Although these promises brought a warm welcome, the program budgeteers soon encountered difficulties. In order to carry out their highly advertised benefit-cost analyses, it was necessary to assess the benefits of various educational programs in terms of

dollars so that these benefits could be compared with costs incurred in achieving them. But no one in education land was willing and able to place dollar values upon the benefits to individuals and to society of various educational accomplishments.

It was during this period of frantic search for a way to assess the benefits of education that our brave budgeteers met another crusading group, known as the behavioral objectivists. Unlike the program budgeteers who were mostly economists, the behavioral objectivists were experts in educational measurement. They believed that if a child learns, his behavior will change and it is possible to measure the amount of the change by suitable tests. From this simple proposition they concluded that educational goals and objectives should be stated as expected changes in student behavior and that teaching effectiveness should be measured by changes in student behavior.

Although the program budgeteers and behavioral objectivists started their reform movements separately, it was inevitable that they should discover each other and find that they had much in common. The program budgeteers thought they had at last discovered someone who could evaluate the educational product and put a dollar value upon it so they could make their benefit-cost analyses. The behavioral objectivists, on the other hand, were not especially interested in the mysteries of the budgetary process, but they liked the idea of getting a little bit closer to the source of money. Besides, by forming a union, both crusading groups were better able to protect themselves from the hostile humanists whose attacks were growing in number and intensity every day. So the two reform groups worked together, sharing the same platform and seeking converts to their common cause.

After a short period of cooperation, disagreements began to emerge. The program budgeteers seemed obsessed with costs and the amount of resources needed for various programs and goals. They were interested in the budgetary process culminating in the allocation of resources to various educational programs. The behavioral objectivists, however, were interested in instructional planning and

evaluation, and their activities culminated in the selection of a teaching procedure.

Both the program budgeteers and behavioral objectivists found it impossible to place dollar values upon the behavioral objectives which had been formulated. For example, they were unable to determine the value to society of teaching all children selected aspects of Roman history. Moreover, they were unable to determine the cost of achieving various behavioral objectives because the cost depended upon the methods used and the ease with which the students learn.

It became increasingly clear that there are important distinctions as well as interrelationships between (1) a management system designed to allocate resources among various educational programs and (2) an instructional planning and evaluation system designed to utilize educational resources most effectively.

Moreover, instructional planning and evaluation are part of the ongoing teaching process and therefore should continue to function after the budgetary decisions have been made for a fiscal period. Clearly, the two systems function on different time schedules. The budgetary process is necessarily geared to the fiscal year while the instructional planning and evaluation process is, or should be, continuous.

Another distinction stems from the divergent concerns of the two systems. The budgetary process is primarily concerned with decisions that affect costs such as class size, salary schedules, the maintenance of summer schools, and so forth, whereas the instructional planning and evaluation process is concerned primarily with finding more effective teaching procedures and techniques which often have no effort upon annual costs.

These distinctions make it increasingly clear that the program budgeteers and the behavioral objectivists should cooperate, but they should remain sufficiently independent so that each can make its unique contribution to education.

The Factory Analogy

One of the inherent weaknesses of the articulate educator is his tendency to promise more than he can deliver. Fifty-two

years ago the authors of the seven cardinal principles of education accepted as a school's responsibilities such goals as: health, worthy use of leisure time, ethical character, worthy home membership, and civic responsibility as well as command of the fundamental processes. In response to these glowing promises school attendance and school budgets increased substantially, yet we seem to be no closer to these goals than we were in 1918. We promised more than we were able to deliver and since the public expected more, its disappointment has been great.

The behavioral objectivists concur with this diagnosis, pointing out that if we had used behavioral objectives in 1918 when we wrote the seven cardinal objectives we wouldn't be in this trouble today. This is perhaps true, but there is also a danger of promising more than can be delivered in student performance. When students are regarded as "products" of the school, it is implied that the school is a factory and should be fully accountable for the behavior of its products. Yet, without the power to select its raw material and reject defective products, the school cannot guarantee its product.

To regard an 18-year old youth in the graduating class as a product of his high school implies more responsibility than the school can accept. His behavior is influenced by many factors beyond the control of his school. The time-honored allies of the home and school in the child-rearing process -- the church, the YMCA, the Boy Scouts, the Camp Fire Girls -- are still with us, but they are losing their effectiveness. So is Horatio Alger's message to youth that hard work and dedication bring success.

Along with the deterioration of these familar guideposts is the triumphal entry of television. The new Pied Piper is preempting more and more of the waking hours of children and youth, providing a powerful distraction from the school's three R's, which have low Nielsen Ratings. Unfortunately, this twentieth century Pied Piper spends a disproportionate amount of time selling soap and cereals.

With all of these factors contributing either positively or negatively to the education of children and youth, to call them

the products of the school is to promise more than the school is able to deliver. The factory analogy leads inevitably to a demand for a quality control system in which all enrollees of the school are expected to perform at a preestablished level of skill as measured by appropriate tests. The great diversity of human talent and varying contributions of the home and peer groups makes this degree of uniformity impossible.

If the factory analogy and the guaranteed product are rejected, for what should the schools then be held accountable? The answer, it seems to me, is quite clear. The school should be held accountable for the scope and quality of the educational services it renders to its students. For example, a school may well guarantee the following educational services for its students: (1) No first-grade pupil shall be in a class with more than 25 pupils. (2) All pupils whose reading scores are more than one year below their grade norm shall receive 15 minutes of individual tutoring each day. (3) Summer classes shall be available for any student who failed a high school course required for graduation. Other statements describing the scope and quality of educational services could be added.

These statements emphasize the educational services to be available to the student. For these services the school can and should be held accountable. Moreover, the effectiveness of these services in contributing to student test performance and to subsequent successful living should be under constant review. But even with such checks upon the school's effectiveness, it is impossible to guarantee the test score of an individual student in the same way that a manufacturer guarantees his product. For these reasons, then, the school must be accountable for the scope and quality of educational services it renders to students even though it cannot guarantee an individual student's test performance as implied by the factory analogy.

It will be noted that although the scope and quality of educational services rendered to students is an output of the school system, not an input, it does provide essential information needed for estimating program costs. For this reason, the identification

//

of the scope and quality of services to be rendered is especially relevant to the budgetary process. By way of contrast, the behavioral objectives lack cost relevance. It is not possible to determine the cost of teaching a child the multiplication tables. The cost is a function of the method used to achieve the behavioral objective, not of the objective itself. Thus, while the evaluation process appropriately emphasizes behavioral objectives, the budgetary process must be concerned with the scope and quality of the educational services to be rendered to students.

Benefit-Cost Relationships

The inherent difficulties encountered in placing a dollar value upon behavioral objectives make benefit-cost analysis in the true sense of the word difficult, if not impossible, in the field of education. However, significant cost comparisons can be made if either the goals or the amount of resources used are held constant.

For example, two different ways to achieve the same educational objective might be tried. If the same goal is achieved by both methods, a strong argument can be made for accepting the less costly of these methods. In this analysis alternative ways to achieve established educational goals are compared. Notice that a dollar value to society of the goals has not been established.

Another kind of analysis is possible in which the amount of resources expended is held constant but these resources are used in different ways to achieve unequal educational outcomes. In this analysis, the educational procedure which achieved the greatest educational output would, of course, be preferable. Such cost effectiveness studies can be made because they do not call for placing a dollar value upon the educational benefit itself.

In a very real sense these two analyses indicate a fundamental change that is occurring in American education at this time. When I first started teaching, back in the early thirties, the superintendent of schools made speeches to the chamber of commerce about getting the most for the school dollar. He didn't talk about PPBS or about benefit-cost relationships, but he did talk about getting the most for the school dollar. His approach reflected the

prevailing notion that the needs of education are unlimited but the number of dollars which society can allocate to education are limited. He wanted to get as much money as possible for the schools and then to spend it wisely, maximizing the amount of educational service rendered.

The basic assumption underlying his approach to budgeting -- that the needs of education are unlimited -- is being questioned today. Under the concepts of both the program budgeteers and the behavioral objectivists, limited educational goals are established first, then the least costly way to achieve them is sought. In the past, it was assumed that the goals of education were indeed unlimited and the problem was to get the most education for the available school dollars. We are now being asked to specify rather precisely limited objectives for education and then search for the most economical way to achieve them. In one sense both procedures are the same, since they both seek the most cost-effective use of school dollars. Their differences are more subtle.

Under the unlimited educational goal concept, boards of education and professional educators determine educational priorities after the legislature and the taxpayers have determined the amount of funds to be available. Under the limited educational goal concept, the legislature and the taxpayers review and approve the educational goals and then provide sufficient funds to attain them.

Along with these changes in the locus of the goal-setting function, hard-pressed taxpayers are challenging the unlimited goal doctrine. Specifically, they are asking: "How can we tell when too much is being spent on public education?"

This is a question that educators usually have avoided. When an answer was given, it often emphasized unmet educational needs. Such an answer is no longer acceptable since there are unmet needs in every phase of American life. To answer the question, the benefits of another million dollars for public schools must be compared with the benefits of an additional million dollars for welfare, for defense, for pollution control, and so on.

This leads inevitably to value judgments which are always involved when important educational issues are resolved. And this is

where the humanists are likely to ambush the program budgeteers and behavioral objectivists. The ensuing conflict will pit the superior statistical skills of the program budgeteers and behavioral objectivists against superior rhetorical skills of the humanists. The outcome is in doubt.

Regretably, this encounter will be based upon the assumption that PPBS and behavioral objectives are, per se, good or bad for education. The truth is, they can be either. The right kind of each in the right quantities could be just the tonic our ailing school system needs; an overdose of either could be fatal.

ISSUES IN IMPLEMENTATION

I

Mark R. Shedd

I would like to begin by stating the obvious. Accountability in education, particularly in public education, is long overdue: No longer can anyone involved in the education of our young people, whether it be the superintendent of schools, the district superintendent, the principal, the classroom teacher or the teacher's aide, be classified as a sacred cow--above reproach and immune from having to prove to the public that he is doing his job and doing it well.

Accountability may be a tough word for many educators to swallow, but I'm convinced that the actual digestion of the accountability process in education will be much less painful than the first bitter taste. Of course, there is still a question or two to be answered; such as, what <u>is</u> "the accountability process?"

To answer that question, I'd like to take a brief look backward at the educational scene since the passage of the Elementary and Secondary Education Act of 1965.

Since then, we have as a nation poured billions of dollars into compensatory education throughout the country. We have looked to the results of those efforts and we have found better schools, more libraries, more paraprofessional personnel, higher paid teachers and administrators, and higher tax burdens.

Yet, in the face of all of these improvements, there is no major study that shows at this time any significant improvement in pupil learning, reduced dropout rates or better teaching. Unlike professionals in medicine we are not able to say that our clients measurably are better for the effort.

One of the few things we can prove conclusively is that education can consume large sums of money with few measurable achievements to show for it.

This record, particularly as it is seen through the eyes of those outside education, raises some pretty basic questions of accountability. For instance, Governor Russell Peterson, of Delaware, speaking before the Education Commission of the States last July in Denver, put it very bluntly when he said:

"Educators traditionally think in terms of inputs--new programs, more dollars for materials, higher teacher salaries and the like. We have files, and wastebaskets, full of statistics about education--how many schools, how many teachers, how many strikes and campus rebellions.

"In principle, the American educational commitment has been that every child should have an adequate education, but when a child fails to learn, school personnel have all too often labeled him 'slow,' 'unmotivated' or 'retarded.' Our schools must assume the commitment that every child shall learn...."

Nolan Estes, the superintendent in Dallas, speaking before one of Educational Testing Service's previous accountability conferences, described his own situation by saying:

"Our target had been those schools in which students were averaging only a half-year's achievement gain for every full scholastic year. By the time we finished, we had not managed to improve on this sad record; in fact, some of our Title I schools were worse off in 1970 than they had been in 1965.

"Five years and five billion dollars after Title I was passed," he added, "We still have not learned how to break the

...derachievement that sees children from poor homes do ...ol; find poor jobs or none; marry--and then send ...r children to school."

...stark fact remains that, after years of experi- ...er spending millions upon millions of dollars, and ...e of the most sophisticated technological and ...in the world to the problem, we continue to grind out ...f functional illiterates from the educational mill ...ear.

It is certainly no wonder that the public does not believe it is getting its money's worth from public education and has begun to beat the drums for accountability.

Yet, there is no magic "accountability" button that one can push, nor is there any one segment of the public educational structure that can bear the brunt of accountability.

Traditionally, as Fred Hechinger, education editor of the New York Times points out, it has been the student who has been held accountable for his performance in school. "The classroom," Hechinger contends, "has been the teacher's castle, and schools have largely escaped the burden of responsibility for any lack of student achievement."

Then, with the advent of massive federal remedial education programs in the mid-'60s, the finger began to point at the home as not holding up its share of educational accountability.

Now, as we embark on the decade of the '70s, it has become the "in" thing to do to make the teacher the whipping boy for the breakdown in the educational process. The public, angered by escalating educational costs, primarily due to large increases in pay each year for teachers, has grabbed at the concept of accountability and is attempting to use it as a whip against teachers.

This I object to very strongly. Accountability should be a process by which we can show ourselves how to teach better and how to turn out a better educated child, not as a weapon to beat a dead horse.

Accountability will get nowhere, as it most certainly should not, if it is cast in the punitive sense of saying to a teacher: "Here, you SOB, raise all these kids' achievement levels a certain rate or you'll get fired."

Accountability must involve everyone from the superintendent of schools to the classroom aide. It must also involve the parent and business and industry as well. And it must involve processes as well as people.

Terrel Bell, USOE deputy commissioner for school systems, put the whole matter of accountability in proper perspective when he told a recent ETS conference:

"Accountability looks at school resource deployment, materials selection, time allocations, and a host of other school management practices. Needless to say, accountability has many facets, forms, and faces. It reaches far beyond the simplistic assertion that it is concerned with teachers and teaching. When the students fail to learn, the entire system must be introspective."

With this in mind, a look at the experiment in New York City involving ETS and the Board of Education is most heartening. Briefly, the New York Board has contracted with ETS to come up with an accountability design to define performance objectives for both students and staff, with the ultimate goal of recommending an administrative structure for an accountability system.

What is heartening is the assurance of everyone concerned that the system will be a positive rather than negative one, aimed not at threatening job security, but at making it possible for everyone to do a better job.

With that assurance, the New York teachers union, the United Federation of Teachers, has backed the plan. Union President Albert Shanker joined in the announcement with the Board, proclaiming that the accountability system would "provide the greatest protection teachers have ever known."

He explained this by saying that successful teachers would be protected against unfair criticism by providing proof of their

effectiveness, while the system would indicate the additional training needed by ineffective teachers.

I would certainly consider this a workable agreement, one which would help dispel the fear among professionals of the word, "accountability."

Yet, it <u>is</u> ironic that militant teacher organizations who had spoken so strongly against accountability formulas until the New York experiment, may themselves have been one of the reasons for the accountability clamor.

Traditionally the public has viewed teachers and school adminstrators as practicing professionals, and has placed a high degree of trust in the professional judgment of school personnel. That, however, was before collective bargaining and teacher strikes.

It was not long before the public, conditioned to the fact that teachers had chosen to deal with their employers in the same way employees negotiated with large industrial concerns, began to ask: "What do we get in return?"--in much the same way that corporate management asks how many more or better cars or television sets will be produced as a result of the contract settlement.

Paradoxically, school personnel are insisting that the public continue to trust them as professionals in the same way that physicians and attorneys are trusted. Just as an example, few patients would think of trying to negotiate a guaranteed cure with a doctor. Rather, we expect the doctor to exercise skill and professional judgment in treating our ills.

Educators have, however, negotiated a different managerial ethic than that enjoyed by other professionals. It may be a big step in the right direction, simply because the outcome of the negotiations has been accountability.

But I think I've talked enough about the reasons for accountability. I'd like now to get, briefly, to three of the accountability situations that have developed in Philadelphia. Two involve guaranteed performance contracts, about which I won't

say very much because I'm sure other speakers at this conference will delve deeply into this aspect of accountability. The third involves our own home-grown brand of accountability applied to a district-wide reading program.

For the record, we do have the nation's largest guaranteed performance contract under which the Behavioral Research Laboratories, in return for $600,000, has guaranteed to raise the reading achievement levels of 15,000 pupils by one grade level in one year's time on a money back basis for every pupil who doesn't make it. The tests have already been given, and we'll know in a few short weeks how we have done.

The other performance contract is with the Westinghouse Learning Corporation, through the nationwide, 20-school district USOE project, involving 600 students in reading and math. Here, too, we'll know shortly where we stand.

But what I'd really like to talk with you about is our eight-district reading program, involving all 285,000 students in grades K-12 throughout the city. In assessing our difficulties with teaching Johnny to read, it became painfully evident that we were sadly lacking in an adequate delivery system--a system that would assure a plan's success by following a tough, predetermined set of goals and objectives based on a process of accountability.

We began to devise this system, in conjunction with the personnel of our eight school districts, by planning eight different attacks on the reading problem, each one geared to meet the specific needs of the children in that district. It did not come easily.

First, we had to gear up our research capabilities to measure quality by output, rather than input. Educational research going back to the '20s always measured such things as expenditure per pupil, the number of library books, dollars spent on teachers' salaries, and numbers and statistics almost down to the number of wash bowls per student.

Relatively few, if any, educational dollars were spent on charting the output, the actual achievement of the pupils. This we did through the application of standardized tests, and the results in Philadelphia, as in all the other major cities across the nation, were devastating.

To put it bluntly, 40 percent of our students were functionally illiterate.

Our first function was to develop and have approved by the Board of Education specific, no-nonsense, system-wide reading goals that would make us accountable to the pupils, their parents, and the public for reaching certain new plateaus of reading achievement in a given time period.

This we did, and the goals are bold ones. I'd like to quote from the board resolution which set them up.

"By the end of the school year 1975-76, each student receiving instruction in regular classes will acquire those reading skills necessary to adequately function in school and society; and students in Philadelphia will read as well or better than students elsewhere.

"It is further resolved that the following indicators will be accepted as evidence that the goal has been accomplished:

1) All students who leave the Philadelphia schools at age 16 or older shall read on the 8th grade level or above; and

2) The scores of all students on an appropriate test shall indicate that the percentage of students who read at grade level or above is greater than the national norm; and

3) All children completing the sixth grade shall read at the 5.5 level or above."

Admittedly, in light of past failures, these goals will be extremely difficult to meet, to say the least. But we honestly feel we can do it.

Once the goals were announced, we were faced immediately with the task of cranking up a management system capable of attaining them. In effect, we simply had to build a better mousetrap.

We did, in fact, build a new delivery system, and we hope it's a better one. It consists of a technical task group of reading experts whose job it will be to ride herd on the entire reading improvement program; a detailed statement of system-management needs to be followed; alternative implementation plans to throw into the breech until we find one that works; and projections of time, costs, and resource demands.

But, of course, building this type of management delivery system left us with two more problems: We had to develop new skills and management capabilities within existing school system personnel through staff and leadership development programs; and we had to recruit top-level management persons from outside the school system, particularly in previously neglected areas such as data processing, planning, budgeting, finance and research. This, too, we have accomplished.

Finally, after undertaking all this extremely complex retooling, we had to tie it all together if we expected it to work at all. The thread we used was accountability.

We were firmly convinced that accountability in a large urban school system trying to teach children how to read was no more difficult than accountability in a large industry trying to reach its own production goals. If business can hold its employees accountable for production, and attain success through that accountability process, then a school system should be able to do the same thing.

Our accountability is a two-way street.

The Board and the Superintendent not only are accountable to the public for reaching their predetermined reading goals, but they are also accountable to their employees to deliver to them the supplies, equipment and administrative support and know-how to accomplish the task.

In turn, each district superintendent is accountable for support to his reading project manager, his principals, and his

teachers, as well as being accountable to the Superintendent for producing results in his district.

The principals are accountable to their teachers and to the district superintendent; and the teachers, of course, are accountable not only to the principals, but also to the youngsters they are teaching.

It must be understood clearly that the teachers rely on district and central office delivery of goods and services in order to carry out their tasks. Administrators in charge of delivering these components are, therefore, accountable to the teachers.

We hope, obviously, that with all this accountability built into a new, considerably more sophisticated management delivery system, we can, at long last, produce concrete results in the form of measurable gains in student achievement and attitudes.

We have, in essence, established a hard-to-reach goal, handed each individual district the ball, supplied them with the very best coaching possible, and told them to run like hell.

And run they have. The districts have elected to use more than 40 different approaches to reading instruction, consistent with the needs of their individual youngsters.

Besides involving BRL in the guaranteed performance contract, the districts are working with McGraw-Hill, SRA, Reader's Digest, EDL, Macmillan, Lippincott, Addison-Wesley, Random House, and many more, on various ways to teach reading to children with various needs.

The firms are supplying materials and resources right from preschool through high school to meet the reading needs of the districts. Methods range from the traditional to the boldly experimental, from multi-lingual programs to computers, from pipelines into universities to self-contained units.

In other words, we have tried to put it all together, but on a step by step, district by district, school by school, child by child basis. We are using project management techniques,

delivery systems, data processing, and just about everything we could get our hands on--all tied in with accountability.

We agree, as you do, that youngsters should, indeed, have the right to read. And we agree further that, although proper teacher preparation for, and performance in, the inner city school will continue to be critical measures in determining whether we reach our goal, we feel strongly that school management will have to face up to its share of the burden, get up off its posterior, and crank up some programs that aren't doomed by their very makeup before they even get to the classroom.

For this, we must be held accountable.

Yet, through all the birth pains of accountability, and its attendant pressures, I have a fear, a strong fear, that marks will take precedence over the kids themselves. James Cass, education editor of Saturday Review, sums it up extremely well when he writes:

"As we focus increasingly on pupil performance as a measure of teacher-effectiveness, it would be easy to forget the complexity of the learning process--that individual children are very different, that they learn different things at different rates, and that even the same child learns at a different rate at different times.

"If, therefore, the laudable effort to improve classroom practice by assessing teacher and school effectiveness merely results in placing more intense and sophisticated pressure on the children to perform, the very principle will be denied in practice, for if the concept means anything, it is that the ultimate accountability must be to the children."

To that I would add, "Amen." And to you I would say, thank you. It's been a pleasure to address you.

-30-

I S S U E S I N I M P L E M E N T A T I O N

II

Francis Keppel

In speaking to this topic at other ETS conferences, the late Robert W. Locke of McGraw-Hill used as a subtitle, "Accountability Yes, Performance Contracting, Maybe." As his replacement on this occasion, I can both honor his memory and try to carry out his work by quoting from his remarks about performance contracting:

> Contrary to what you have read in the papers, I believe that most education companies look upon performance contracting as an undesirable way of doing business.
>
> For companies with carefully researched programs and the competence to train teachers, it is not particularly risky because they know what kinds of results they can achieve. However, it puts them in a straight jacket that makes performance contracting less desirable than the same work done under a simpler contract. The reason why certain companies, such as my own, have responded to the recent surge of RFP's is simply that they have the programs, they can provide the services, and they are willing to take the risk in order to get the business. A customer is a customer.
>
> It is worth noting that many large and well-run companies have <u>not</u> sought to win performance contracts, either because they consider the risk too high or simply because they have reservations about their ability to perform the requisite services.

The companies that <u>are</u> willing to make performance contracts -- and perhaps all education companies -- would surely agree on these things:

1. That results in education cannot be guaranteed. In the fall and winter in Texarkana there were some misguided claims about programs that could "produce grade-level independent readers and writers by the end of the first grade" or some such, but virtually all companies understand that intellectual processes cannot be guaranteed in the way that soapmakers guarantee cleanliness.

2. They understand also the critical importance of having reliable data around which to construct contracts. More on this point later, but the lack of sufficient data is probably the main reason why some responsible companies have been reluctant to make performance contracts.

3. They also recognize that performance contracts make more sense for innovative programs than for conventional ones. It is hardly worthwhile for either school districts or the companies to write such involved agreements merely for the purchase of textbooks. It makes much more sense for the installation of complex new systems of instruction for which the learning environment will have to be reorganized and the teaching staff retrained. This may be the chief value of performance contracting, because innovative instructional programs are very difficult to install and yet hold much promise for the improvement of education.

I suspect, in short, that the education companies have much the same general views of performance contracting as the school systems that wish to hire them.

I believe that Bob Locke accurately reported on the concerns, the hopes, and the uncertainties of those considering performance contracting. Let me ask you to look at two issues that seem to me to lie behind our attitudes on this aspect of "accountability." The first is related to public policy and the adoption of programs by public bodies. The second has to do with the state of the art of establishing educational objectives and measuring results. To what extent can public authorities and the private sector rely on our present knowledge and techniques?

There are, of course, a good many other issues that could be discussed in this context. Consider, for example, the probable impact in the decade 1968-1978 of the fact that enrollment in the elementary schools will <u>decrease</u> by five percent at the same time that the Office of Education estimates a steady <u>increase</u> in costs. Educators and educational companies will find themselves in the same boat: legislatures and school boards will want some new explications for rising costs if the old argument of rising numbers can no longer be used. The best argument may be that rising costs bring better processes and better quality of learning in the schools. The mere fact that school authorities are trying out new ways of doing things may be a factor which may be helpful in persuading legislatures almost regardless of results. But it may also mean that performance contracts will be in trouble if the legislators and the public feel that they involve "high" profits, even though the profit might be quite commensurate with that accepted for other businesses of comparable risk.

Or consider another fact: that today the schools devote a very small part of their annual budgets to purchasing materials and services directly related to instructional programs from outside the system. One can argue about what should be included in this category, but I doubt if the percentage figure is likely to exceed 2.5 percent per annum, or something like $25 per pupil per year on the average, which is to be compared to average per-pupil expenditures for all purposes in 1968 of $696. In considering the wisdom of wide-scale performance contracting, therefore, responsible authorities in education will have to take into account the capability of the

so-called educational industry to undertake, on any large scale, nationwide contracting which would involve a large increase in that percentage. For the truth is that the educational companies are not big business today. They do not have a large staff of experts in the management of educational programs, or of specialists in the various aspects of schooling. Few if any of them would be able, or perhaps willing, to mount large-scale, geographically dispersed, and high quality programs in a hurry, particularly in view of the possibility of relatively low profit margins. This is especially true, I think, when one considers the two factors of public policy and the state of the art to which I now ask you to direct your attention.

Questions of Public Policy and Legal Authority

We have a deeply rooted tradition that the state has the obligation and the duty to provide an educational program at public expense. Though the state is the sovereign power in regard to schools, much policy-making power has been delegated to the school boards elected locally. When policy-making functions are delegated to public officials or bodies, they may not be redelegated to a private entity. In all phases of performance contracting from the contents of the Request for Proposals to the contract itself, the board must reserve the policy-making decision to itself and delegate only "ministerial" functions. The dividing line between these categories is obviously difficult to draw, and so far as I know there are few judicial determinations to help us with this specific subject. There are many details to be considered, varying from state to state depending upon the state constitutions, statutes, rules, and regulations. As long as the contracts are experimental in nature or are to procure consulting services for a district, there appears to be no legal obstacle that cannot be overcome.

Performance contracts let to date -- ranging from the Gary contract for taking over an entire school to contracts merely to improve the skills of selected students in selected subjects -- have not, so far as I know, yet been challenged in the courts. If performance contracts become a standard practice, however, they surely will threaten the interests of some group who well might

resort to litigation. Your legal advisors and counsel for potential contractors can thread their ways through the legal problems and give appropriate advice so that performance contracts -- of one sort or another -- can be legally entered into; but whether such contracts will meet the objectives of the districts entering into them remains to be seen.

One likely possibility is that the public contracting authority, in order to carry out its policy-making functions, will have to specify in the contract the procedures, personnel standards, curriculums, and relations for the social environment of the school so precisely that the performance contractor will in his "ministerial" functions have little room to move around. State laws and judicial interpretations may limit the range within which innovative programs can be conducted. The irony of the situation is at once apparent: Performance contracts entered into with the intent of creating change and taking advantage of flexible operations may end up inflexible. This may not be what the school system had in mind as an objective. And from the point of view of the performance contractor it obviously reduces the attractiveness of the arrangement from both the professional and the profitable point of view. Considered as a long-range business possibility, tightly controlled contracts (in the sense of limitation on personnel and permissible procedures), which may be subject to annual negotiation or cancellation, are not necessarily appealing, especially to a small-scale industry which would have to spread itself across thousands of school districts in the United States.

Questions Related to the State of the Art

How well performance can actually be measured is a technical issue, and a complex and serious one at that. Standardized tests, like any other measuring devices, are subject to error of measurement. Indeed, it is a far greater relative error of measurement than that usually found in the engineering field from which performance contracting has been borrowed. One consequence of this large error of measurement is that individuals who have been chosen because they scored low on a pretest have a very good chance of scoring impressively

higher on a correlated posttest even if little real learning has taken place. In the absence of true experimental control groups or longitudinal measurement, regression toward the mean may work to the advantage of the performance contractor, if not the student.

There is strong pressure currently to use criterion-referenced tests in place of standardized norm-referenced tests as measures of performance and achievement. These have the attractiveness of comparing a student's performance with a seemingly objective standard, instead of with the performance of other students. They also are likely to be more specifically tailored to the objectives and content of particular units of instruction than are most standardized tests. But there is no guarantee that a criterion-referenced test is any less subject to skill sampling errors than standardized tests, and eventually the standard of satisfactory performance on a criterion-referenced test depends on someone's judgment of what is acceptable or desirable. Who shall make the judgment? Certainly not the performance contractor.

I do not for a moment intend that this brief excursion into a few of the technical intricacies and uncertainties of testing obscure a very real need. Our educational system has been inequitable in its distribution of resources to develop human potential. It has acted selectively on racial and economic bases. Hopefully one thrust of accountability is the intention to redress that selectiveness. I would hope that in the process we do everything possible to assure that our objectives, procedures, and test scores are what they purport to be: genuine credentials to educational, social, and economic mobility, and not temporary passports to a later, more tragic letdown.

My conclusion on the technical issue at this point is that the standardized tests must be used at some point -- they're the best we have at the moment -- but that the criteria of satisfactory performance in a performance contract need the most careful explication, and probably should include follow-up measures beyond the life of a project or program.

The issue of transferability gets to the heart of the functions of performance contracting. The turnkey function is the

one that we hear much about, and the one that in many respects makes the most intuitive sense. It includes elements of installation, demonstration, and training. The contractor thus is apt to be (or should be) particularly interested in the contract planning process. Who is involved? How did they become involved? What is the long-range plan? What action is being taken that gives some assurance that the school intends to commit itself to adopting the new program if it is successful?

As far as I know, the turnkey function of performance contracting has yet to be shown to occur. And, in my opinion, if it does occur it will be to a large extent unpredictable, given the present state of our knowledge and techniques. At the very least it will be contingent upon success of the contract, and results of even the current wave are not in yet.

I am not trying to lead to the proposition that I see no prospects for transferability. The point I want to make is that it is extremely important for both the school and the contractor jointly to think through the role performance contracts are really supposed to play in the plans and programs of the school, and to develop means of providing a continuity of good results. The point indeed leads to another issue, that of feasibility.

There are questions of feasibility from the viewpoint of the contractor, the school, and both together.

Feasibility for the contractor is, I think, largely a matter of his personnel capabilities, and of how much he knows about his product (books, machines, syllabi, techniques, and so on) -- what they can do, with whom, and under what conditions. And this will of course depend in large part on how well they have been tried out and evaluated. Trite as it sounds, if the contractor does not have extensive and reliable performance data in his files, he is courting nasty surprises. Both school and contractor need to review evidence of expected performance carefully and in good faith before entering into an agreement.

I think of feasibility from the point of view of the school largely in future terms. Obviously the school has little to lose

fiscally although it may suffer a time loss with a performance contract if the project doesn't work. The questions are: Does the school have the ability to follow through with successful projects? Does it have the planning know-how? Does it understand what it may take to adopt and support the program?

Part of this future aspect of feasibility concerns the time period of the contractor's responsibility. For major remedial programs, which I take the current larger contracts to be, a few years are probably enough to show that immediate performance results can be obtained. But it is not enough to show that they can be maintained. And it may well not be enough to explicate why they came about, despite the contractor's brochures and the school's report to the public. I think both school and contractor have a responsibility to see that there is sufficient evaluation to make it reasonably clear what the important reasons for success most likely are, and one necessary condition for that is time. Yet school boards may not legally be able to contract for the necessary long time period because of their inability to bind their successors in office.

Some Tentative Conclusions

1. The dividing line between policy making and ministerial functions is now fuzzy. If and as it becomes more sharply defined, there is a possibility that performance contracts may lose the advantage of the very flexibility of operation which is argued to be their greatest strength.

2. The education industry today is neither large enough nor capable of sufficiently widespread geographical dispersion to mount a large-scale program of performance contracts. Relatively few companies have the combined capabilities and experience to handle programs involving teacher training, materials development, evaluation, and information systems. Nor are they likely to develop such capabilities on a large scale if the return on the investment involved is likely to be both low and uncertain.

3. The state of the art of setting educational objectives and of measuring results is insufficiently developed to assure responsible public authorities that they can at once turn with confidence to a massive use of performance contracting to solve their problems.

4. Present practice and legal restrictions make long-term contracts difficult to achieve. Yet the state of the art would seem to require such length of time before public authorities can be confident that they are acting in the public interest.

From these considerations I can only conclude that performance contracting is still in the research and development stage, and should be so considered by both public authorities and the education industry. Too much should not be promised, nor should the idea be discarded too hastily for reasons of current educational politics. We have a long time ahead to sort out the aspects of public education that can safely and wisely be entrusted to the private sector, and those that can not. The state of the art counsels caution, and the state of public education suggests restraint on rhetoric. As Bob Locke put it, "Accountability, Yes. Performance Contracting, Maybe."

P U B L I C E X P E C T A T I O N S

James E. Allen, Jr.

ACCOUNTABILITY -- WHO PUSHES?

Ladies and Gentlemen -- welcome to the Chicago stop of the Accountability Bandwagon!

What I want to know is, are we here just as barkers at the latest medicine show or are we here as real physicians, seeking a new cure?

Are we just making big talk to sell the same old elixir that will soothe and comfort but won't really help or change, or are we seriously brewing a new wonder drug?

Certainly, the ills and problems of education today require something more than a mere soothing syrup, and the real question is: Are we just trying to make the schools look good or really be good?

The answer to this question lies, it seems to me in another question: Why accountability now?

As a matter of logic, accountability should always have been a major factor in education -- and of course, it has not been ignored -- but too often it has been hit or miss in operation and woefully unscientific in design and application.

Is its present popularity a response to outside criticism and lagging support, or is it a response to an inner sense of failure? If it is just the former, we are wasting our time and

might as well go home and forget the whole thing, for such a basis for accountability can all too easily distort its purpose and limit its usefulness.

Certainly our school systems want to win back lost confidence and gain more support, and accountability is the best and surest way to accomplish this. But such a result will be attained only if accountability has as its aim genuine improvement and the elimination of deficiencies that can stand as justification of confidence and increased support.

Unfortunately, in the minds of too many, accountability has much too much of a public relations orientation, and this could be fatal -- destroying the potential of one of the most hopeful developments in education, and making it just another of those cyclical fads that seem to sweep through the educational world, causing great commotion but leaving things little changed after their passing.

Accountability is more often than not considered in the narrow sense of assessment and measurement rather than in its broader and more definitive meaning of being responsible and liable. Thus defined, the question becomes: responsible to whom?

If we accept the premise that the schools belong to the people -- the premise upon which our public school system has been developed -- then the answer is obviously the public. But what is the public?

Is it society as a whole? It has generally been accepted that the education system serves all of society by producing valuable members, and this service encompasses everything from training people to be good citizens, to teaching them useful skills, to making it possible for each individual to reach self-fulfillment.

Is it the taxpayers? Here the responsibility seems very clear. The taxpayer supports the schools and deserves the assurance that his investment is worthwhile. The concern here generally is that the schools be run efficiently and that the tax bill be not too great.

Is it the parents? They look to the schools to educate their children, and by this they usually mean that they expect their children to be able to participate effectively in society and to be prepared to make a living.

Is it the students? This is the most direct and the most difficult responsibility, for the school must deal with the dual task of considering and satisfying the needs and interests of both the present student and the future adult.

Obviously, the concept of accountability as responsibility is a very complex one, and the schools cannot be considered as accountable in any simple sense to any one group. But whatever part of the whole may be in mind, the essence of concern is not with methods and means of accountability but rather with the matter of responsiveness. Certainly, then, crucial in the objectives of accountability is an increased responsiveness of the educational system to those whom it is supposed to serve.

Who then influences the system? To what pressures do the schools respond?

Students, parents, and taxpayers have the influence of consumer and provider. As consumers, there is little opportunity beyond individual and organized protest and persuasion to influence policy. As providers, the most effective action is the negative one of withholding funds, with too little opportunity to distinguish between reasons of tax burden and dissatisfaction with school performance.

The education profession as a whole is of course enormously influential. Teachers organization contracts contain provisions about many things beyond salary and working conditions. The profession itself for the most part governs the accrediting agencies. Education lobbies exert great influence on state and federal legislation. Schools of education prepare the teachers and administrators who run the schools; they prepare also many of those in state and federal educational agencies. Organizations such as the AFT, NEA, and AASA have nationwide constituencies and wield tremendous political power.

Local boards of education have the responsibility for the immediate operation of the schools, and their policy decisions, though shaped within the limits of state and federal laws and policies and relationships to the profession, in large measure determine the nature of local educational opportunity.

The state has the legal responsibility and authority for the provision of education, and its fiscal, regulatory, and supervisory functions profoundly affect the quality of local performance.

The federal government because of the nature and scope of present conditions and trends in education is assuming a much more direct and influential role.

Among all these interests bearing upon the schools, seeking to influence their performance, where is the push that gets action? Where is the force that can make accountability a working reality?

In my judgment, it is in government and the profession.

Even with better means of exerting their influence and wielding their power, the public can have only a limited role. This is not to underestimate the importance of the public's role. Indeed, the public has to become more aggressive and unrelenting in its insistence upon performance as a measure of quality.

A promising note for the future of stronger and more active public concern for education is the increasing interest of youth in working for change within the system and in gaining greater support. It was highly encouraging to me to have a student at Princeton consult me last week on the feasibility of organizing a national youth movement in support of education. What better group to reinforce the drive for accountability! I strongly encourage such a movement and urge all concerned youth to take part. Not only can the young bring the insights of immediate experience, but they can also serve as a constant reminder that it is the product and not the process that must be our constant concern.

But even with the maximum exercise of their power, the public can only deserve and demand action -- government and the profession have to provide it, and the real push must come from these sources.

It cannot be just gentle pushing or polite nudging. For too long a kind of protective pussyfooting, both in government and the profession, has characterized efforts to bring about real change in education. In seeking to avoid stepping on sensitive toes, we have too often succeeded in protecting the profession at the expense of the children and youth -- a true silent majority if there ever was one, just now beginning to raise its voice and assert its claims.

In a broad sense, the educational responsibilities of government at the state and federal levels and those of the profession divide, with much interacting and overlapping, into that of the responsibility of government for creating those conditions in which good education can flourish and that of the profession for producing good education within those conditions.

The ability of state and federal governments to create conditions favorable for good education will be greatly strengthened by three actions which should be immediately gotten underway. The first is a complete overhaul of the patterns of school finance involving at the federal level a combination of consolidated categorical aids and of broadly based general aid, allocated on an equalization basis, and, at the state level, assumption by the state of all, or substantially all, of the local costs of elementary and secondary education.

The second deals with structure. At the federal level education should be elevated to Department status with a Cabinet-level Secretary and a national advisory council.

Department rank would ensure the sharper focus of greater visibility and higher status, and an advisory council would provide the opportunity for an advocacy of education's concerns, unhampered by undue influence of partisan politics or special interests.

The council should be bipartisan with its members appointed by the President for staggered terms of sufficient length to prevent excessive dominance by any one President or political party. Its basic role should be advisory, not administrative. As an established body, not tied to the four-year election cycle, it could provide a continuity that could help to mitigate the disruptions of changes in Administration. It could enlighten the nation by making an annual report to the President and the Congress on the state of education. It could act as a conscience for the people of the country by both reflecting and arousing concern for the well-being of education.

At the state level, a strengthening of structure is essential. As the states are now organized and managed, they are, in varying degrees, unprepared to encompass the expanded dimensions of their educational task or to deal with the broadened concepts that now define their role. Thus, each state must examine its own capabilities and undertake, to whatever degree may be necessary, an overhaul of its arrangements for the governance of education. Expanded federal support for education should incorporate incentives for modernization of state education structures.

The third action required to aid government in its responsibility for creating the conditions for good education is a systematic program of research and development. This aim would be substantially furthered by Congressional enactment of the legislation establishing a National Institute of Education, proposed last year by the President.

Obviously an adequate discussion of these actions would be a speech in itself, but I cite them because they are essential to full effectiveness of the governmental role in accountability.

Despite the widespread interest in accountability, I doubt that it will advance on any large scale to the action stage without the determined leadership and constructive, focalized help of state and federal governments in both ways and means.

Government must assist in developing and making available the technique and instruments of accountability. Government financing of education must recognize the need for accountability and use the power of the purse as leverage to encourage -- indeed, where necessary, to require -- accountability.

Within the profession there is growing support for accountability, but as is too often the case, it comes primarily from those who are already productive, with support lacking where accountability is most needed.

The resistance to accountability is understandable. In the minds of many, it sets up an image of excessive testing and measurement, inimical to initiative and imagination. It implies criticism, generates defensiveness, and leads to a kind of pass-the-buck attitude.

An acceptance of accountability requires a perspective that recognizes the indivisibility of education and concentrates on its purpose.

This kind of perspective has become increasingly difficult as the numbers to be dealt with, and the expanding dimensions and character of education, have produced an undue concentration upon the parts rather than the whole.

Despite the vast machinery of the system, however, education's purpose is still simple -- the development of the intellect and the discovery and encouragement of abilities and talents. In attempting to achieve this purpose on such a large scale, each of the various parts of the enterprise has itself become so immense as to demand a degree of attention that excludes, or makes very difficult, a constant mindfulness of its place within the whole.

This kind of compartmentalization and particularization of focus, this concentration on narrow self-interests, must be discouraged for they act to destroy the concerted approach of all components of the system that is required for reform and change -- for acceptance and adaptation to a concept as basic and far-reaching -- and hopeful -- as accountability.

In relation to this point, I read with much interest the New York Times report of a speech by Buckminster Fuller, delivered last week at a meeting of the American Association of Museums, in which he said: "We have so many specialized abilities we can blow ourselves to pieces, but we have no ability to coordinate ourselves. I see our society as very powerfully conditioned by its reflexes, with very, very tight ways of functioning. And that is dangerous -- so dangerous that if man does not stop thinking locally and make the grade as a world man we may not be able to continue on this planet." The news report continued -- "Citing the need for 'synergism,' thinking in terms of the whole rather than its parts, Mr. Fuller urged the development of more Leonardo-type men, who could think in such terms."

More synergistic thinking within the profession can be a strong aid to a widespread acceptance of the necessity of accountability.

By this time, if there was ever any doubt in your minds, it must be obvious that I believe the accountability bandwagon is on the move because accountability is the most promising cure for many of education's most serious ills.

The public to whom accountability is due is becoming more and more aware of this possible remedy and is also much more sophisticated and able to detect any attempts to substitute more of the same old brew in new bottles.

The responsibility for prescribing this remedy and for getting it into action rests primarily with government and the profession -- and if we fail now to follow through, to make accountability a reality, the resulting loss of confidence and withholding of support will make the present doubts about the effectiveness of our schools look like a veritable avalanche of approbation.

THE ROLE OF EVALUATION IN ACCOUNTABILITY

Henry S. Dyer

AND VICE VERSA

Three events in the history of American education illuminate some of the more important roles that evaluation must play in any system of educational accountability.

I

The first event occurred in 1647 when the Great and General Court of the Massachusetts Bay Colony enacted what the history books refer to as the Old Deluder Satan Law.[1] This, you will remember, was a law that sought to foil the designs of the devil by insisting that every child in the Colony be taught to read and write. It held each town accountable for providing this instruction out of its own funds. And it backed up its mandate with an annual fine of five pounds to be levied on any town that failed to comply.

One reason, no doubt, that the Puritan Fathers were able to get away with this high-handed infringement on local autonomy was that there was general agreement in those days on the ends and means of education. All children must be taught to read so that they could have direct access to the Scriptures and thereby have an outside chance of avoiding eternal damnation. One of the major problems in education these days is that people are not all that clear and convinced about the ends and means of education.

One reason for this state of affairs has been suggested by Lawrence Cremin:

> "...too few educational leaders in the United States are genuinely preoccupied with educational issues because they have no clear ideas about education... They have too often been managers, facilitators, politicians in the narrow sense. They have been concerned with building buildings, balancing budgets, and pacifying parents, but they have not been prepared to spark a public debate about the ends and means of education."[2]

Another reason for the fuzziness about ends and means is that educational goals as commonly formulated by educational philosophers have tended to be cast in such sweeping generalities and remote ideals that they have left school people at a loss to use them meaningfully for assessing the actual on-going operations of their institutions. This statement is not intended to denigrate the efforts of educational philosophers. Their ideas are a necessary, if neglected, ingredient of the process by which usable goals can be defined and applied in concrete instances. But they are only the beginning of the process; the gulf between the expression of educational ideals and any practical measure of their realization is so wide and deep that few if any working educators have been able to find their way across it.

The educational oratory speaks of goals like "self-fulfillment," "responsible citizenship," and "vocational effectiveness;" the assessment of school efficiency in specific cases usually depends on such measures as retention rate, college-going rate, average daily attendance, and performance on reading tests. Whether there are any rational connections between the numbers and the slogans is a matter that is rarely considered. The assumption seems to be implicit, for instance, that the longer a youngster stays in school, the greater will be his chances of self-fulfillment, or that the higher his reading score, the more likely that he will become a responsible citizen. But such assumptions are left largely unexamined, and in particular cases may be obviously wrong. In short,

the answer to the all-important question, "Accountable for what?" is left hanging in mid-air.

Therefore, one important decisive role that evaluation must play these days in any educational accountability system, which is not designed solely to find scapegoats to assuage our collective guilt, is that of helping all of us sort out and evaluate our educational goals and objectives, so that we can begin to get some definite and agreed-upon ideas of where we want the schools to be taking us as well as our children, and what we think the priorities ought to be.[3]

Over the years there have been some promising efforts in coping with this problem of goal-setting at a practical level, but a lot still remains to be done if the community served by the schools is to become as deeply and significantly involved in the process as it must be if the notion of accountability is to make any sense at all in shaping education to fit the individual needs of the pupils as well as the needs of the troubled society that they are going to inherit.

II

The next historical event, illustrative of another aspect of the accountability doctrine in education, occurred nearly 300 years after the enactment of the Old Deluder Satan Law -- in 1930 to be exact. This is a bit of personal history, for it was in 1930 that I had my own first traumatic experience of being held professionally accountable as a teacher. I was in my first job teaching senior English. I had one particularly weak student whose parents were bound and determined that he should be shoe-horned into a certain prestige college that I firmly believed was well beyond his capabilities. My principal gave me to understand in no uncertain terms that, for my part in this process, I was to be held accountable for seeing to it that the boy passed the old style College Board exam in English at a level that would make him admissible to the college his parents had chosen for him. The implication was that if the boy failed to make it, the renewal of my contract would be in doubt. In short, my performance as a teacher was to be evaluated, at least in part, on how that student performed on that exam.

Back in the '30's the College Board exams -- unlike those of today -- had passing scores which were defined in terms of performance criteria laid down by the examiners. Today, I suppose those old-fashioned exams of forty years ago, with all their presumed faults, would have probably qualified as "criterion-referenced tests." It is curious how history -- even in testing -- seems to be repeating itself.

In any case, what did I do to prove my accountability in that situation? How did I go about getting students to meet the criterion set up by that old-time criterion-referenced test in English? I did what many other high school teachers were doing in those days. I crammed my students on all the old College Board exam questions of the preceding ten years, filled the kids up with canned themes so that they might appear to write profoundly, though possibly a bit irrelevantly, on any topic that the examiners might dream up, and ground the standard literary classics into their heads until they were thoroughly sick of them.

By so doing I fulfilled my obligation and my contract was renewed. My weakest student passed the English entrance exam with flying colors. He was admitted in September 1930 to the college his parents had chosen. He flunked all of his mid-semester examinations in November 1930, and was fired shortly thereafter. By meeting my obligation under the narrow definition of teacher accountability then prevailing I had succeeded in preparing the student to become a failure in college.

What does this episode suggest about the role of evaluation in an accountability system? It suggests that if the system is to work to the benefit rather than the detriment of the young people who go to school, we must be continually observing and evaluating the side-effects and the after-effects of what goes on in classrooms. For if, by the processes we employ, we teach children to pass tests at the expense of learning to hate the subject in which we test them; or to hate the whole idea of learning, it seems to me we defeat the whole purpose of education and fail to be accountable to the students themselves.

The armamentarium of educational and psychological measurement contains a good many instruments of various types for evaluating students' attitudes toward learning, toward themselves, and toward one another. Admittedly these instruments are still pretty crude. The state of the art in the measurement of attitudes, values, and the like was summed up by David Krathwohl and his collaborators in these words in their book on educational objectives in the affective domain:

> "...we cite many techniques for appraising such objectives, but we are fully aware of the fact that much must be done before the development of testing techniques in the affective domain will reach the rather high state of clarity and precision which is now possible in the cognitive domain."[4]

Nonetheless, if, as we have been saying all along, the schools are to be concerned about the developmert of the whole child, we had better make judicious but regular use of the best of these techniques. Be it noted, however, that such techniques should not, in my view, be used as a basis for evaluating the children themselves. They should be used, rather, as a basis for coming as close as possible to evaluating the full impact that schooling may be having upon the lives of the children. Insofar as schools fail to do their best to seek out this kind of evaluative information about themselves regularly and routinely they are failing to be accountable in any educationally acceptable sense of the word.

III

The third historic date in the development of the principle of accountability in education was April 1965 -- the date when the Elementary and Secondary Education Act was signed into law. You will recall that clauses 5 and 6 in Section 205 (a) of the original Act provided that procedures be adopted for annually evaluating programs designed to meet the needs of educationally deprived children and that the evaluative data accruing from these

procedures were to be incorporated in annual reports from each local education agency to the state education agency and thence to the federal government.[5] The purpose of course was to try to account for the incremental educational benefits that the federal dollars were buying, and it is this aspect of the evaluation/accountability equation that is understandably uppermost today in the minds of many taxpayers and their representatives on school boards and in legislative bodies.

In view of the agonizing fiscal crises in so many school districts, this is, of course, a legitimate and urgent concern. It is a concern, however, that generally has overlooked the difficult problem of providing the needed evaluative information. The authors of one intensive study of the early functioning of ESEA have said that "when ESEA was in its first weeks and months of implementation ... the infrastructure of <u>systematic</u> program evaluation was either monexistent or woefully primitive."[6] Anyone who has kept up with attempts to evaluate ESEA programs -- particularly Title I programs -- in the last six years knows that this statement is still largely true, in spite of some noble efforts to lick the problem. At least part of the reason is that there are still nowhere near enough people out there in the school districts who know how to put a dependable and meaningful evaluation program together -- one that is capable of genuinely and dependably relating educational benefits to educational costs, and this despite numerous attempts to apply to the educational enterprise such appealing notions as cost-effectiveness, planning-programming-budgeting, management information systems, and the like.

Be all this as it may, it seems to me that the most important aspect of Section 205 of ESEA is <u>not</u> that it appeared to hold local school systems accountable for making educational expenditures produce a measurable pay-off in pupil learning. In point of fact it did nothing of the kind. If you read the original Act carefully, you come to realize that all it called for was merely a <u>rendering of an accounting</u> -- an evaluation, if you will -- of what was going on in Title I programs and how well they were working. The big emphasis was, and still is, on objective and accurate annual

reports on how the educational process is functioning on behalf of students and how much money is being spent in the effort. This is a type of annual report that had never been produced before and, to my knowledge, has not been produced yet. We may know how much we spend on textbooks, on teachers' salaries, on busing, on food service, and so on, but we still do not know how to cost out a program in elementary school reading, or high school science, or health, or whatever in such a way that we can actually isolate the costs of each program per se and relate those costs to the children's growth in reading competence, or their love of books, or their physical well-being.

When you put the problem of rendering an accounting in this way, you may well begin to wonder whether the problem, like that of squaring the circle, can ever be solved. It suggests that, in approaching the question of how to render an accounting of what is going on in an educational system, there is a real question of how far the accountability concepts that may be useful in the control of industrial systems can be applied to school systems. For the production of learning and human development is hardly analagous to the production of soap or cat food or space vehicles.

Moreover, the measurement problem in each case is just about as different as it can be. In trying to achieve accurate measurement of the inputs and outputs of the industrial enterprise, one is concerned with making the human factors in the measurement process as small as possible, and in many areas the instrumentation for this purpose has become remarkably automatic and efficient. In the measurement of the cognitive and psychosocial functioning of students, however, the human factors are the very essence of what we are trying to measure and evaluate. Consequently, when we speak of measuring such human qualities as problem solving in mathematics, or teacher effectiveness, or vocational aspirations, we are speaking of a process that is vastly different from that of measuring electric power output, or the noise level in communication lines, or the trajectory of a missile. Indeed the difference is so great that an atomic scientist concerned with measuring the speed of electrons once

suggested to me that we should probably drop the word measurement altogether when dealing with educational and psychological phenomena.

He may well have been right, for I suspect that much of the misinterpretation and over-interpretation of test-score data that bedevils so much educational thinking stems from the failure to realize that the metaphor of the yardstick, or the chronometer, or the ammeter, or whatever is a wholly inappropriate metaphor when one is trying to evaluate pupils' development and the educational programs and environmental conditions that affect it.

I do not intend the foregoing to mean that, in some appropriate sense, the measurement of pupil performance is a hopeless or futile endeavor. Quite the contrary! Furthermore, such measurement is indispensable if we ever expect to render a rational rather than a purely intuitive accounting of how schools and school systems are doing. But the rendering of such accounts in education is not likely to be very sound or instructive if educational decision-makers think that assessing the quality of human learning and development is on all fours with measuring the quality of widgets.

IV

To recapitulate briefly at this point, what sort of perspective on the evaluation/accountability equation do the three bits of history provide? First, the Puritan Fathers who wrote the Old Deluder Law were so sure of their educational objectives and the means by which they were to be attained that they were able to get away with holding every school district accountable for providing a particular type of instructional service. They did not, however, concern themselves with the evaluation of the effects of the instructional service provided, since they assumed that that would be taken care of by more remote means on the Day of the Last Judgment. They were apparently unaware of the possibilities of evaluation as a form of self-correcting feedback.

Back in 1930, I was held accountable for producing a certain single measurable result, and by that result my performance was evaluated. There was, however, no obligation upon me to account

for the means by which I obtained the result. The feedback was sure and swift, but it was what Norbert Wiener would have called defective feedback inasmuch as it included no information on any side-effects or after-effects my teaching methods may have been having on the student.[7]

ESEA holds school districts accountable for rendering an accounting -- that is, for providing an evaluation -- of the effects of the programs being federally funded, but it says nothing about any punitive action that might be taken if the hoped-for results of the programs are not forthcoming. That is, it calls for effective evaluative feedback -- which incidentally it has not yet been able to get in any comprehensive way -- but it does not specify how the feedback would be used if it were obtainable.

In looking back over these three aspects of the role of evaluation in the evaluation/accountability equation, one gets the feeling that something is missing and that that something is to be supplied by a reversal of roles. In addition to thinking of the role of evaluation in an accountability system, one needs to think also of the role of accountability in an evaluation system. Which is to say that if educational evaluation programs are to serve any useful educational purpose, then those who support and manage school systems must be made accountable in three ways: (1) for seeing to it that the evaluative information the programs provide is as good as it can be, (2) for seeing to it that the information is interpreted within the limits imposed by the nature of the data, and (3) for seeing to it that the information is <u>used</u> in some systematic fashion to find ways of continually bettering the quality of instruction for all the children in all the schools.

A final comment or two on each of these three points is now in order.

1. <u>How to make sure that the information an evaluation program provides is as good as it can be</u>. This means first of all selecting tests and other instruments that are well-crafted and well-validated for the purposes to which they are to be put. There is a considerable body of literature on how to make such selections

and an even larger body of measures from which to select.[8] This material should be conscientiously examined before picking any test for use in the schools. Second, it means that the tests shall be administered in a manner that guarantees, insofar as possible, that the students know what they are expected to do and that they will do the best they can. This may seem painfully obvious but the fact of the matter is that test data are too often invalidated right at the source because of maladministration. Finally, and equally obviously, the tests must be scored with scrupulous accuracy. I mention these humdrum rules only because I am impressed by the fact that the failure to observe them is usually overlooked as a possible explanation of why the pupils in some schools appear to perform surprisingly higher or lower than their counterparts in other schools.

2. <u>How to make sure that the results are interpreted within the limits imposed by the nature of the data</u>. Here we are in considerably deeper trouble because it is abundantly clear that most consumers of test results seem to be amazingly unaware of the limitations of such data. One of the glaring problems in this connection is that of getting those who make educational decisions on the basis of test scores to realize that the best of achievement tests is never more than a sample of a student's performance and is therefore inevitably subject to sampling error. This simply means that if his score on, say, an arithmetic test places him among the bottom third of his classmates today, his score on an alternate form of the same arithmetic test taken the very next day has a good chance of placing him among the middle third of his classmates.[9] Failure to recognize this inherent bounciness of test scores can and does lead to all sorts of mistaken conclusions about the effectiveness of remedial programs for students who are selected for such programs on the basis of their low achievement test scores.

Another glaring problem in the interpretation of academic achievement tests has to do with the kinds of numbers in which the measures are customarily expressed -- namely, so-called grade-equivalency scores. Except for the notorious IQ, these are probably

the most convenient devices ever invented to lead people into misinterpretations of students' test results. Both the IQ and grade-equivalency scores are psychological and statistical monstrosities. I have elsewhere defined the IQ as "a dubious normative score... that is based upon an impossible assumption about the equivalence of human experience and the opportunity to learn."[10] A grade-equivalency score has many of the same properties, and as such it lures educational practitioners to succumb to what Alfred North Whitehead called "the fallacy of misplaced concreteness."[11]

There is not enough time here to go into all the irrationalities that underlie the construction of grade-equivalency scales, nor all the misconceptions they generate in the public mind about what achievement tests are saying about how well students and schools are doing. Instead, I urge you to read a recent brilliant paper by Roger Lennon, entitled "Accountability and Performance Contracting."[12] Lennon's credentials are among the best, since he is senior vice president of the company[13] that published two of the most widely used and well-constructed achievement test batteries -- the Stanford and the Metropolitan -- both of which, interestingly enough, have grade-equivalency scales attached to them. I have said the paper is brilliant; one might also call it courageous, because in it Lennon, from his own intimate knowledge of the subject, spells out in considerable detail much that is absurd, wrong, and misleading about grade-equivalency scales and why such numbers should <u>not</u> be employed for expressing the results of otherwise useful achievement tests, or for assessing professional accountability, or for determining how much educational contractors should be paid.

In his frank discussion of this and other similar problems in the interpretation of educational measurements, Lennon nicely exemplifies an important aspect of the role of accountability in educational evaluation.

3. <u>Finally, how to use evaluative data in a systematic fashion to find ways of continually bettering instruction for all the children in all the schools</u>. This, it seems to me, is the major task that lies ahead, if educational evaluation is to fulfill its promise.

And it brings me to the questions in your conference program that I am expected to answer. I shall now try to answer them:

1. Can the relevant inputs, outputs, and conditions of operation [of educational systems] be satisfactorily measured? The answer is, "Yes, for the most part they can be, if school systems will make the kind of informed and serious effort required."

2. If so, what are the appropriate techniques? The answer to this question is necessarily somewhat complicated. But it does have an answer, and anyone who is seriously interested in finding it might begin by examining two of the articles that appeared in the Phi Delta Kappan of December 1970 -- the one by Stephen Barro and the one by me.[14]

3. If not, what remains to be done? I have already said that adequate evaluative techniques are available if one has the will to use them. Nevertheless, it must be said that we do need better measures than we now have of the personal-social development of students, better measures than we now have of the many factors inside and outside the school that influence students' over-all development, and more particularly better ways of observing and describing what actually goes on day by day in the teaching-learning process. By this I mean that we need far better ways of systematically monitoring and describing what is really going on behind the facade of fancy labels by which we characterize so many so-called innovative programs like I.T.A., I.P.I., G.S.A., the Open Classroom, the Discovery Method and so on ad infinitum. I am convinced that we can obtain these kinds of information if we have the will to do so.

4. Finally, are different techniques needed for different types of educational systems? And here my answer is, "Yes, but..." Yes, the evaluative techniques one would use for a small homogeneous educational system, such as one might find in a suburban community, would be different but also less satisfactory than those one would use for a large heterogeneous system, such as one might find in a large city. But the best way for small homogeneous systems to secure the most useful evaluative data about the effectiveness of their educational programs is for them to join forces, for evaluative

purposes, with other systems, possibly on a state or regional basis, so as to enhance the possibility of uncovering, through well-worked-out statistical analyses involving all the schools, those educational innovations that have the best chance of paying off for their own students.

My answers to all of these questions are meant to imply that an evaluation system expressly designed to keep the quality of instruction continually rising will be a highly complex system. One might prefer something simpler. But I suggest that, in the highly complex world in which we now have to live, simplistic approaches are not likely to help us much in finding our way to education for either the good life or the good society.

REFERENCES

1. Records of the Company of Massachusetts Bay. II. P. 203.
2. Cremin, L. A. The genius of American education. New York: Random House, Vintage Books, 1966. P. 111.
3. See for instance: Stake, R. E., "Objectives, priorities, and other judgment data." Review of educational research, 40:2 (April 1970). Pp. 181-212.
4. Krathwohl, D. R., and others. Taxonomy of educational objectives: Handbook II: Affective Domain. New York: David McKay, 1964. P. 87.
5. Public Law 89-10. 89th Congress, H.R. 2362. Sec. 205(a), (5) and (6).
6. Bailey, S. K., & Mosher, E. K. ESEA: The Office of Education administers a law. Syracuse, New York: Syracuse University Press, 1968. P. 162.
7. Wiener, N. Cybernetics; or control and communication in the animal and the machine. Cambridge, Mass.: M.I.T. Press, 1961.
8. For example: Buros, O.K. (Ed.). The sixth mental measurements yearbook. Highland Park, N. J.: The Gryphon Press, 1965; also Educational Testing Service.
9. Dyer, H. S. "How Precise Can Measurement Be?" in Evaluation and Christian Education (Spaulding, H. F., Ed.), New York: National Council of Churches, 1960. Pp. 44-54.
10. Dyer, H. S. "The Functions of Testing -- Old and New" in (Tabor, Victor, Ed.) Testing Responsibilities and Opportunities of State Education Department, (undated). P. 67.
11. Whitehead, A. N. Science and the modern world. New York: Macmillan, 1944. Pp. 74-75.
12. Invited address, AERA Convention, February 6, 1971.
13. The company is Harcourt Brace Jovanovich of New York.

14. Barro, S. M. "An Approach to Developing Accountability Measures for the Public Schools"; Dyer, Henry S. "Toward Objective Criteria of Professional Accountability in the Schools of New York City." Phi Delta Kappan, 52:4 (December 1970).

T H E F U T U R E O F A C C O U N T A B I L I T Y

Edythe J. Gaines

A WAY OUT OF THE TANGLED WEB

O what a tangled web we weave when first we practise to -- to change what educational consumers (pupils, parents, public) had in mind when they applied the term accountability to education into something professionals in education believe they can live with more comfortably. Rather simple-to-understand and straightforward notions have been all but buried in a mass of calculations, complications, protestations, and concatenations. Taking that route, our profession is in danger of missing a grand opportunity to save the profession, to raise it to new and higher levels of respectability and of status, to bridge the ever-widening gap between educators and their various publics, and to gain a much-needed measure of peace across the battle-scarred fields. These in recent times have tended to separate us from the newest claimants of our services, that is, the "new" minority groups in this country, one of which, Afro-Americans, is at least four hundred and seventy-eight years "new" inasmuch as they arrived with Columbus.

Let me remind you that the question of accountability was not raised, in the first instance, by the professionals. It was raised by the laity. It tended to be raised in three contexts.

First, in communities where citizens vote directly on education taxes, the question has been raised in the context of school bond issues. "Parents and taxpayers" groups, with very strong emphasis on "taxpayers," have rebelled against the rising costs of education without receiving an equivalent rise in client satisfaction with what the education dollar is buying.

Second, when legislators were persuaded to vote allocations of funds especially for programs designed to improve educational results for the "economically and educationally disadvantaged" student, they insisted upon proofs, to be provided by evaluation studies, that the money spent brought about the results intended.

Third, there were parents and students and members of the public who were the traditional supporters of education -- usually supporting bond issues, higher education taxes, higher teacher salaries, and greater expenditures for education generally -- who began to have serious doubts as to the ability of the public schools to educate their children effectively. The patently-easy-to-see, clearly evident massive failure of the schools to educate the children of the poor and of minority groups brought cries for accountability from these groups early on. Upon closer inspection, parents from other walks of life began to perceive that their children, too, were not getting the kind of education to which the parents aspired. Coalitions began to form -- shaky but there, nonetheless -- between and among a broad spectrum of parents, all demanding better performance on the part of the schools and couching their demands in terms of professional accountability. While this group traditionally has not begrudged education a full measure of financial support (quite the contrary), today it, too, is asking searching questions about the uses to which educational resources are being put, in terms of the results that are forthcoming.

It is with reference to this last-mentioned group that my remarks will deal. It is this group whose alienation should be of particular concern to public school people, for it constitutes our traditional base of support. Loss of this support base leaves us

with no foundation at all on which to rest. The continued existence of public education hangs in the balance. Can anyone deny that?

Given the stakes, one must be appalled at the response our profession has made to calls for accountability on the part of this group of people. When parents asked why their children were not getting the kind of results from schooling they had hoped for and expected, the response has been one or more of the following.

One response: There is something wrong with your child. The message is never given in so straightforward a manner as this. It is couched in such descriptive terms as culturally deprived, socially disoriented, linguistically handicapped, educationally disadvantaged, neurologically damaged, genetically weak, perceptually impaired, emotionally unstable. But these terms are almost never accompanied by unassailable evidence about the accuracy of the label as applied to that individual child. These terms tend, instead, to be group labels, and parents rightly are not about to accept such unproven labels as being applicable to their child. Moreover, parents know that something must be wrong with the diagnosis when too many children fall within the "problem" categories. Parents easily accept the thought that one or two percent of the children may fit one or more of these descriptions. When the figure gets to be more than 10 percent, parents begin to be concerned about the accuracy of the data. Although they may not complain, they become alert and watchful. But when the "problem" or failing children get to be 60 percent to 80 percent of the total pupil population, as is the case in several poorer neighborhoods in the city, then nothing is going to convince parents that the problem lies with the children and not with the schools -- and they are right!

Another response: There is something wrong with the surrounding environment. Again, other terms are used -- such as social disorganization, societal factors, the poverty cycle. The problem with this is that parents know all about the surrounding environment (after all it is they, not the professionals, who live there). But they fail to see, and the professionals have failed to prove, a

one-to-one relationship between a child's reading retardation and whom the child's parent is sleeping with. Indeed, if sleeping only with one's bona fide, certified husband or wife were a precondition for learning to read, then precious few children from the upper middle to upper upper social strata would learn to read. Parents are mindful of the fact that the public schools in the case of the earlier minorities, the European immigrants, specifically accepted the responsibility of bringing the children of the poor into the mainstream of life in this country. They do not understand why the schools have lost that earlier sense of mission and the accompanying skills to carry it out.

Another response: The schools cannot be solely responsible for all of today's ills, including pupil failure; the schools cannot be held accountable for everything. But, dear colleagues, don't you know that parents already know that; that parents are fully aware that schools cannot be expected to do everything. However, the distance between not being held accountable for everything and not being held accountable for anything is quite a distance indeed -- one that parents and students simply cannot be expected to accept. We can't have it both ways, you know. We can't say to the public on the one hand that the schools are institutions vital to the general health and welfare of this nation, thus deserving strong and expensive support, while saying on the other hand that the schools cannot assure the nation that they can bring about any of the outcomes expected of institutions which are vital to the general health and welfare of the nation.

Another response: Standardized test scores cannot appropriately be used to judge the effectiveness of individual teachers nor of school units. But think a bit. Why do parents place such confidence in the value of standardized test scores as measures of success and effectiveness? The answer is that we, the professionals, have taught them to so regard these scores. Look at how we have used them. We have deemed them to be so accurate a measure that we use them to group children; to select those who are going to get "enriched"

curriculums and those who are going to get educational pablum; to determine who will go into the "academic" track and who will go into "commercial" or "general" tracks in high school; to determine who will go to college and who will not. Indeed, we have used them to determine the life chances of our youth.

Given this example on the part of professionals, is it any wonder that parents have logically concluded that standardized test scores are quite accurate and quite legitimate measures of success and effectiveness? Why not, then, apply these miraculous measuring tools to a determination of the success and effectiveness of individual educators and of school units?

No, it is we educators, not the laity, who are dealing in non sequiturs and logical inconsistencies when we say (through our behavior if not our words) that these test scores can make clear determinations abut the performance of pupils but not about educators and schools.

Again, we cannot have it both ways. Parents undoubtedly will give up the idea of rating educators in terms of test results when educators give up rating children by these measures. Only when poor scores do not automatically denote a poor student (and vice versa) will parents concede that poor results do not automatically denote poor educators.

Another response: The concept of accountability leads to the measurement of the narrowest aspects of education, those aspects that can be quantified and measured. Yet, the most important outcomes of education are human and humane, and these will not yield to an accountability scheme. Therefore let us not have one. Methinks thou dost protest too much. Parents and students have never advocated accountability measures limited to numerically quantifiable characteristics. Quite the contrary. They are very much concerned with organizational climate, with attitudes and nonverbal communication, with the quality of the human transactions that occur in schools. Clearly, too, they want much more from schools than whatever skills are measured by standardized tests and other such tools. I repeat

that such tools are much more the darlings of the professionals than they are of the laity. We created them; we gave them status; we use them pervasively.

No, parents are not satisfied merely with accountability with respect to the "3 Rs." Parents simply say that schools ought at the very least be able to teach the "3 Rs" after centuries of continuous experience. It is precisely for this reason that parents are appalled to see that schools cannot even succeed in teaching some of their pupils mastery of these ancient and basic tools for further learning.

Another response: Since certain children are doing poorly on standardized tests, let's not administer these tests to them at all. Constant failure on these tests is destructive of the children's confidence and sense of self-worth. Let us not subject them to this unproductive pain. Indeed, let us choose to teach other things at which these children can be more successful. I must confess that the first time I heard this argument I couldn't believe my ears. Yet, I've heard this argument advanced again and again and by people who consider themselves to be "on the side of these children." (Please note the expression, "these children.")

Parents hardly know what to say to this argument. It is so chock full of patronizing insult to their children. It is a clear call for goal displacement.

This is not to say that standardized tests are good or valid or well-constructed or anything else. It is to say that parents reject a double standard in education. They also suspect that behind this "kind" argument lies a desire on the part of educators to hide the evidence of their failure to effectively educate "these children."

Another response: Accountability is an extremely complex notion. It requires knowledges and tools not presently available to us. We do not have adequate measuring tools. In fact we don't even have an accurate and agreed-upon definition of that which is to be measured; therefore, we cannot really have an accountability system at this point in time. Here we get a circular argument. When all

the limitations and restraints and complications are explained to them, parents say that they were never asking for an accountability system measured with the accuracy of calipers. They say that they'd be content to have effectiveness bear some relationship to results as seen in pupil performance. But when parents say that, educators immediately warn that student performance is a function of "countless other variables which are often uncontrollable and too multi-dimensional to analyze effectively." The author of that statement (Allan C. Ornstein in a paper called "Methods for Conducting Teacher Behavior Research: With Implications for Educating the Disadvantaged," based on his unpublished doctoral dissertation, N.Y.U.) went on to argue that there is no agreement as to what constitutes desirable student behavior; therefore, there is no way to tell whether or not a teacher is or is not getting the desired results. In other words, you can't get there from here.

Parents rejoin, "Fellows, stop making it so complicated. We didn't ask you to calibrate every aspect of our child's functioning. We just asked you to stop producing functional illiterates who not only hate school but also are frequently totally turned off to learning by the time school is finished with them. We ask you to stop changing bright-eyed, alert, ready-and-eager-to-learn kindergarteners into glassy-eyed, surly, and turned-off secondary youth who think of school mostly as a place to 'get in, get through, and get out.'"

Parents know that it is absurd to suggest that we can measure with absolute precision and exactitude the degree to which an individual professional's input results in a specific and identifiable outcome in terms of individual pupil performance, and they know that it is absurd to suggest that we can get such exact measures with respect to the effectiveness of school units. To use this as an argument, however, to suggest that nothing can be learned about the impact of educators and schools on the performance of individual pupils and groups of pupils is to carry the argument far too far. It flies in the face of common sense and of one's sense of truth.

Every parent knows a teacher who seems always to be successful with whatever children are assigned to him. Every parent has a fair notion of which schools are more effective than others. You ask parents to deny their sense of truth when you ask them to deny that such differences exist, can be identified, and probably can be explained in terms of differentiated human input.

Another response: (This is what I call the throw-yourself-on-the-mercy-of-the-court-with-a-plea-for-clemency response.) Yes, you're right. We educators really don't know much about why certain pupils are failing nor about what to do about it. We are failing with "these children," but there is no malice there. We are doing the best we can. We just don't know any better. So, let's devise an accountability system that will not hold us responsible for present performance, but which will show us the way to become better performers in the future. There is something about this "we don't know" thing that bothers me as a professional. I am reminded of a scene in the motion picture Gone with the Wind in which Butterfly McQueen, playing the role of a house slave who got her "privileged position" in the household of the pregnant Scarlet O'Hara by claiming to be expert in delivering babies, cried out in panic, when the crucial moment came, "But, Miss Scarlet, I don' know nuthun' 'bout no babies!" So with too many of today's educators. They hardly cross the threshold before they begin to tell you that they cannot do what they profess (by title and position) to be able to do. We have, for example, the spectacle of the English major licensed to teach English declaring unashamedly that he knows nothing about the teaching of reading -- as if reading were not part of English.

Parents say, "You called yourselves educators, I didn't. I accepted you for what you said you are. You told me that yours is a profession; I didn't tell you that. I merely accepted what you said. You said that as a professional you are entitled to professional salaries, to professional autonomy, to the other rights and privileges of professionals. I did not argue with that. I merely ask that you do what you profess to be able to do. This is no time to tell me that you don't know how!"

It is for this reason that I do not believe that the accountability scheme being worked out in New York City will be at all acceptable to the groups of parents and community people who demanded it. In essence it is an educational accounting system, not an accountability system. Parents will not object to such an accounting system, but I doubt that they will accept it as a substitute for the accountability system for which they asked. They are likely to consider it to be a giant cop-out, a betrayal of their trust. It is my fervent hope that Educational Testing Service, the chief consultant for the project, will radically restructure its design so that it does indeed at least attempt to meet the people's aspirations for professional accountability -- a concept that rests upon the notion of responsibility <u>that can be fixed</u> for the results of the educational enterprise.

The design for New York City speaks of <u>collective</u> responsibility of the staff for <u>knowing</u> as much as it can about the pupils, and of collective responsibility of the staff to <u>use</u> this knowledge, <u>as best it can</u>, to maximize the development of pupils toward defined and agreed-upon pupil performance objectives.

How in heaven's name can a staff be "collectively responsible" for anything? How do you operationalize that concept? What if they "collectively fail"? Then what? Do you ultimately separate them from service "collectively"?

Really, the public is not a collection of fools. They know that such terms have no real meaning. They know that we know that, too. Therefore, they suspect us of acting in bad faith. That is no basis upon which to reestablish public trust in the public schools and in their educators.

<u>A final response: Parents and public cannot be trusted with data on school performance. They would use such data as weapons, engaging in vigilante activities. Therefore, try to avoid giving them such data</u>. No doubt fear of reprisal is the reason behind many of the diversionary maneuvers in the accountability game. But what is the evidence on this matter? There have been so few cases of this kind

of witch-hunting that every case has been headline news. Why should an entire profession cower in unreasoning fear of unlikely consequences?

In a truly useful accountability system, practitioners would learn what they need to know in order to be successful in their work. First they would learn what is expected of them since parents and public would, at long last, have to combine their lengthy, over-inclusive, often conflicting set of expectations into agreed-upon goals. Next, practitioners would learn the extent to which they are or are not meeting those goals. Finally, they could use that information to help them to search out, in a clear and focused way, those practices and programs which hold promise of helping them learn to behave in such a way as to more nearly successfully meet those goals.

In such a setting practitioners could reach a new level of freedom from stress. Make no mistake about it: with or without an accountability system, the educator who is failing in his work knows it and is shattered by that knowledge. Nothing equals the private anguish experienced by such a person. However, since he perceives it to be to his advantage not to admit to his problem, he tends to take on adaptive modes of behavior that are not merely destructive to the ends of education for children, but are also destructive to the practitioner himself -- both personally and professionally.

These, then, are the most typical responses of professionals to the concept of professional accountability. They are very much out-of-phase with parent and public perceptions, aspirations, and expectations.

What, if anything, could get us professionals to change our stance? I believe that a change will come only when we perceive that there is something vital to us in the concept. We cannot be motivated by fear. No one can be. We can be motivated by hope and by the promise of saving solutions to our own pressing problems.

A good accountability system would free us from the stresses and strains inherent in trying to meet myriad, unspecified, often conflicting expectations. It would recognize not only that pupils are uniquely individual human beings but also that educators are, too. Thus, it would free us to find our own personal paths to glory, if such can be found. It would not subject us, therefore, to the anti-human individuality and uniqueness that inheres in the "collective responsibility" notion. Such a notion leads to collective anonymity and conjures up an image of a long, grey line of faceless, nondescript beings, each indistinguishable from the other -- all working, but none knowing the worth of his work, the relevance of his work to the growth and development of other individual human beings.

A good accountability system would, indeed, help to increase our professionalism, raise our profession to higher status, heal the wounds we've suffered in battles with those who should be and traditionally have been our allies, and bring us all enormous personal satisfaction and peace of mind. Then, and only then, would we be out of the tangled web we have woven.